STUDENT AND SCHOOL CORRELATES OF MATHEMATICS ACHIEVEMENT: MODELS OF SCHOOL PERFORMANCE BASED ON PAN-CANADIAN STUDENT ASSESSMENT

John O. Anderson, W. Todd Rogers, Don A. Klinger, Charles Ungerleider, Victor Glickman, & Barry Anderson

This study explored the relationships between student achievement and student, school and home variables from the pan-Canadian assessment program administered by the Council of Ministers of Education Canada (CMEC): the School Achievement Indicators Program (SAIP) Mathematics 2001. The study also evaluated the datasets used in relationship to their utility for statistical modeling of school performance. Student beliefs about mathematics are positively related to mathematics achievement for both age groups and for both domains of mathematics. As students' use of instructional supports (parental assistance with mathematics homework, computers in the mathematics classroom) increases, there is an associated decrease in mathematics scores.

Key words: student achievement, school performance, multilevel modeling.

Dans cet article, les auteurs explorent les relations entre le rendement scolaire et certaines caractéristiques des élèves, des écoles et de l'environnement familial à partir des données du programme d'évaluation pancanadien du Conseil des ministres de l'Éducation (Canada) (CMEC), le Programme d'indicateurs du rendement scolaire (PIRS) 2001 en mathématiques. L'étude évalue également les bases de données au regard de leur pertinence pour la modélisation statistique de la performance scolaire. Les auteurs notent une corrélation directe entre les idées que se font les élèves au sujet des mathématiques et le rendement en mathématiques tant pour les groupes d'âge que pour les domaines des mathématiques. Plus les élèves font appel à du soutien pédagogique (aide des parents pour les devoirs de mathématiques, ordinateurs dans les classes de mathématiques), plus les notes en mathématiques diminuent.

Mots clés : rendement des élèves, performance scolaire, modélisation à multiples niveaux.

Students' performance in public education is an important concern for several reasons, two of which centre on the large scale of expenditures of public funds on education and the beneficial effects education has on the well being of both individual students and society (Hanushek, 2002). A key indicator of educational performance is the achievement of students on centrally administered tests of learning outcomes (Hattie, Jaeger & Bond; 1999; Keeves, 2002; Shore, Madaus & Clarke, 2000). Testing student achievement at provincial, national, and international levels has become a common mandated practice throughout Canada. Implementations of large-scale assessment programs in Canadian schools are typically predicated on the assumption that they will influence the quality of education. For example, Ontario has established an independent testing organization to conduct provincial assessments with the express purpose "to promote greater accountability in Ontario's publicly funded education system and to improve the quality of education" (Education Quality and Accountability Office, 2002, p. 1). However, the ways in which testing and test results are related to system-level characteristics, such as educational quality, are not well understood. Factors in addition to curriculum and instruction, such as student characteristics on entry, teacher and school traits, and home and community characteristics, have been shown to be significantly related to student achievement scores. These relationships vary from one grade to another, one school to another, and one subject area to another (Fitz-Gibbon, 1998; Ma, 2001; Mandeville & Anderson 1987; Rumberger, 1995).

In spite of the apparent complexities associated with schooling and learning outcomes, the communications emanating from large-scale assessment programs take the form of reporting school and provincial averages of simple test scores or the percentages of students classified in different ordered achievement categories. Typically, these reports fail to take into account the contextual relationships of these scores to student, teacher, school, home, or community traits. Instead, the reported results simply show that on average some schools perform better than others in the different skill areas and at different grades. In some public reports (e.g., Cowley, Easton, & Walker, 2001) schools are ranked in terms of student results on these tests (often by aggregating results across subject areas) to describe and monitor system quality. The publication of these

rankings implies that variation in student performance is solely due to school effects.

Because most information related to human endeavours is complex, the underlying context of factors and situations that influence these results must be taken into account to understand educational performance at the class and school levels (Joint Advisory Committee, 1993; Statistics Canada & Council of Ministers of Education Canada [CMEC], 2000). Raudenbush and Willms (1991) point out that, in addition to student variables such as gender, attitudes, and perceptions, class level variables such as composition and size, diversity of instructional practice and teacher experience, as well as school variables such as climate, size, and academic focus should be considered. Willms and Kerckhoff (1995) suggest that the results of large-scale assessments have to be reported and considered at three levels before results are interpreted: the gross level of test score averages; the level of net productivity in which scores are modified to better reflect the variation in the conditions of schooling; the level of inequality of student characteristics on entry (see also Joint Advisory Committee, 1993, Guideline 2, Part IV, Section *B*, p. 20). Much research has been conducted on international and national assessment datasets yielding a wealth of information about patterns of student and school characteristics: the *Programme of International Student Assessment* (Fuchs & Woessmann, 2004; Willms, 2004), the *Primer Estudio Internacional Comparitivo* of Latin American nations (Willms & Somers, 2001), the US *National Education Longitudinal Study* (Dumais, 2002), the US *National Longitudinal Survey of Adolescent Health* (Watt, 2003), state level datasets in Australia (Hill & Crévol, 1999) and the United States (Reeves & Bylund, 2005). However, the national assessment program in Canada, although well over a decade old, has not been well explored and analyzed. The purpose of this study is to explore a pan-Canadian assessment dataset to model student and school traits in relation to achievement and in doing so evaluate the quality and utility of these data for the analyses and modeling of school performance.

The Correlates of Learning Outcomes research program (COLO)[1] is a long-term research program designed to use educational indicators to elucidate the underlying structures and processes of public education

systems in Canada. The results and findings of this research program are of interest and utility to policy makers and educators. The logic of this interest lies in what may be termed the path of policy influence (Kennedy, 1999). The elements of public schooling that are accessible to policy makers (such as funding, curriculum, the nature and extent of instructional support and supervision, the provision of opportunities for professional development, and school organizational structure) should show some influence over the key consequences of schooling (such as student learning, graduation rates, or ease of entry into the labour market). The investigation and understanding of these pathways fits within the realm of evidence-based research and decision making (Raudenbush, 2002; Slavin, 2002) and so, within the scope of the study reported here.

In Canada, indicators of student and school performance and characteristics of students and schools have been collected as part of international, national, and provincial assessment programs. These programs have yielded large datasets containing measures of achievement of significant learning outcomes in the areas of reading, writing, numeracy, and science, as well as indices of student, home, and school characteristics or correlates. Although for Canadian schools where a substantial amount of the score variance can be accounted for by the classroom, the majority of these assessment programs do not permit the analysis of data at the class level (Rogers, Anderson, Klinger, Dawber, Feng, & Zhou, 2006). Consequently, the results reported are restricted to the student and school levels.

The study reported here is based upon the pan-Canadian assessment program administered by the Council of Ministers of Education Canada (CMEC): the School Achievement Indicators Program (SAIP). The assessment examined is that for mathematics conducted in 2001 (see CMEC, 2005 for full documentation) in which mathematics was divided into two achievement domains – Content and Problem Solving. The Content domain included numbers and operations, algebra and functions, measurement and geometry, and data management and statistics. Problem Solving included a range of problems and solutions in the areas of numbers and symbols, ability to reason and construct proofs, providing information and making inferences, pursuing evaluation

strategies, and demonstrating communication skills (CMEC, 2001). The study also evaluated the use of the SAIP data for analyses of this type.

The achievement tests were administered to stratified samples of students, aged 13 and 16, within each province and territory to allow reporting at the provincial and national levels for each of these two age groups. In addition to the achievement tests, the students, teachers, and principals in the sampled schools completed questionnaires generating data on student, class, school, and home variables.

A total of 43,314 students were in the dataset (Table 1) with 1,746 students having no index of which mathematics test they completed (4% loss). This data loss was not considered systematic because it was spread across age groups and mathematics domains, and the level of loss was considered low enough to support the feasibility of further analyses. Issues of data quality, such as missing data, data integrity, and levels of data organization, were encountered and are described separately in another paper in this issue (Rogers et al., 2006).

Table 1: Number of Students in SAIP Mathematics 2001 Dataset

Mathematics Domain	Age		
	13 years old	16 years old	Total
CONTENT	12 334	8 842	21 176
PROBLEM	11 851	8 542	20 393
Missing	1 052	692	1 746
TOTAL	25 237	18 077	**43 314**

THE ANALYSES

The data from the SAIP 2001 assessment of mathematics were divided into four achievement/age groups for analysis purposes: Problem solving – age 13, Problem solving – age 16, Mathematics content – age 13,

Mathematics content – age 16. The analytic approach used was multi-level modeling with the program *Hierarchical Linear* Modeling (HLM5; Raudenbush, Bryk, Cheong, & Congdon, 2000). Student mathematics achievement as measured by the SAIP tests was the level-1 outcome measure and student traits derived from the Student Questionnaire (CMEC, 2005) were the Level-1 explanatory variables. The intercept and gradients of the Level-1 model were then modeled with the school variables (Level-2) that were derived from a principal components analysis of a selection of items from the School Questionnaire and from the aggregation of factor scores derived from the Student Questionnaire.

Analysis of the Student Questionnaire

Analyses were conducted on the 119 items of the student questionnaire to reduce the number of variables at Level-1 into a small set of meaningful factors to make the analysis of the student data more manageable. Further, the student factor scores were aggregated to form school means, producing school-level indicators related to instructional practices, resources, and student attitude.

A set of 32 items was selected from the SAIP Mathematics Student Questionnaire. We used the following criteria: the included items were relevant and accessible to policy making, had relatively good response rates (~10% or less missing data per item) to minimize the effects of missing data, and had relatively high communalities for principal component analysis. Principal component analysis followed by a varimax rotation of the components with eigenvalues greater than or equal to one (Mulaik, 1972) yielded a five-factor solution accounting for 43.3 per cent of total variance. The full student file (n = 43,314) was initially used for this analysis. However when initial analyses were run on the 32 questionnaire items of interest, only 27,677 students (63.9%) had complete data. The missing data was not systematic in that the majority students had complete data records, or at most one or two missing responses to questionnaire items. However, when collated for calculations of correlations, incomplete cases were deleted resulting in the substantial overall decrease in complete records. To address this problem an EM estimation procedure for missing data imputation

(Systat, 2002) was used to estimate and impute all missing data. The final five-factor solution is reported in Table 2.

Table 2: Five Factor Solution: Student Questionnaire

Factor Label *Item stem (paraphrased)*	Questionnaire item #	Loading
Student beliefs about mathematics		
(1=Strongly disagree … 5=Strongly agree)		
Math is more difficult than other school subjects	Q14A	-0.516
I am not very interested in mathematics	Q14B	-0.717
I learn lots of new things in mathematics	Q14C	0.574
Math is an important school subject	Q14D	0.738
Math is important for my future studies	Q14E	0.784
Many good jobs require mathematics	Q14F	0.666
Instructional supports used by students		
(*How often?* 1-Rarely or never 5=Almost every day)		
You & your parents work on mathematics homework	Q20A	0.439
You & your parents work on other homework	Q20B	0.419
In math course we work in pairs or small groups	Q22E	0.386
In math we use math books & magazines	Q23A	0.335
In math we had guest speakers or experts	Q23B	0.672
In math we use computers	Q23C	0.750
In math we use the internet	Q23J	0.771
In math we use the computer lab	Q23K	0.745
Instructional practices		
(*How often?* 1-Rarely or never 5=Almost every day)		
In math courses this year. . .		
The teacher gives notes	*Q22A*	0.428
The teacher shows us how to do problems	Q22B	0.622
We participate in math projects	Q22C	0.391
We are taught different ways to solve problems	Q22D	0.587
The teacher assigns homework	Q22G	0.479
We discuss quiz or tests	Q22H	0.513
We work alone on assigned work	Q22I	0.335

We work on exercises from textbook	Q22J	0.540
We study the textbook	Q22K	0.514
The teacher reads from the textbook	Q22L	0.548
Teachers asks questions of students	Q22M	0.613
Students ask teacher questions	Q22N	0.499
Causes of mathematics performance		
(1=Strongly disagree … 5=Strongly agree)		
To do well in math you need hard work	Q16A3	0.600
To do well in math you need encouragement - teachers	Q16A4	0.789
To do well in math you need encouragement - parents	Q16A5	0.771
To do well in math you need good teaching	Q16A6	0.553
Disciplinary climate		
(*How often?* 1-Rarely or never 5=Almost every day)		
In mathematics courses this year. . .		
There is noise or disorder in the classroom	Q22O	0.792
We lose 5-10 minutes because of disruptions	Q22P	0.784

The first factor was labeled *Student Beliefs about Mathematics* and consisted of six items related to the extent to which students viewed mathematics as useful and important in their lives, and the extent to which the students found mathematics difficult or interesting. High scores on this factor reflect positive student perceptions in the sense of viewing mathematics as useful, interesting, and not difficult. In contrast low scores reflect negative views.

The second factor, *Instructional Supports Used by Students*, contained eight items indicating the extent to which the students used parental help with mathematics and other homework, and the extent to which computers, mathematics literature, and mathematics experts were part of mathematics instruction in their classroom. This factor could be viewed as an index of student use of non-teacher instructional support. High scores on this factor reflect high frequency of student use of these supports, while low scores indicate little or no use of these supports.

The third factor, *Instructional Practice,* consisted of 12 items that dealt with the frequency of various instructional activities in the mathematics

classroom such as working on mathematics projects, asking questions, or working on problem solving. High scores on this factor correspond to a greater frequency of use of a variety of instructional practices in the mathematics classroom suggesting a high diversity in classroom instructional practices. Low scores on this factor could mean that students perceived less frequent use of these practices. Low scores also suggest a more simplified instructional environment.

The fourth factor, *Causes of Mathematics Performance*, consisted of four items related to the attributions students make to mathematics achievement such as working hard, receiving encouragement from teachers and parents, or receiving good teaching. A high score on this factor means that students express a high level of accepting the belief that hard work, good instruction, and encouragement (from both teacher and parent) enhance mathematics performance.

The fifth factor, *Disciplinary Climate*, consisted of the two items on the student questionnaire that dealt with the frequency of disruptions and noise in the mathematics class. Students with high scores on this factor reported that there were high levels of disruption and noise in their mathematics classrooms. A low score on this factor would reflect students perception of low levels of classroom disruption and noise.

Factor scores were calculated for each student and appended to each student's record to serve as student-level correlates of learning. Further, the student scores on each factor were aggregated into school averages that were eventually appended onto each school record for the subsequent HLM analysis. These aggregated factor scores were used as indices of school traits albeit based upon student perceptions.

School Questionnaires

School questionnaires were completed by the school principal in the same time period as the administration of the SAIP mathematics tests to students. A total of 1,681 school questionnaires were included in the dataset. The school questionnaires consisted of a total of 256 items. To make the data set more manageable for statistical modeling, a reduced set of 27 items was selected using the same criteria used to identify the items in the student questionnaire to be included in the analyses: relevance, accessibility, good response rates, and communalities.

The principal components analysis of these 27 items resulted in a five-factor orthogonal solution (Table 3) that accounted for 47.7 per cent of total variance.

The first factor, *Limitations to Learning*, consisted of eight items asking principals to estimate the proportion of single-parent families and students with problems. They also estimated the extent to which various school characteristics limited the capacity of their school to provide effective instruction: concerns about student ability, lack of parental support, community conditions, or shortage of mathematics teachers. A high score on this factor means that a principal perceives that student, home, and community characteristics constrain instructional effectiveness of the school in mathematics. A lower score would be related to principals who view the student and home traits of the school to be less constraining, perhaps even positive elements in enhancing school effectiveness.

The second factor, *Climate*, consisted of four items related to staff morale, school spirit, staff and student pride in school, and community support. A high score on this factor would indicate that the principal views staff, student, and community support and spirit for the school to be positive, and a low score would indicate negative perceptions of spirit and support.

The third factor, *Parents*, consisted of eight items indicating the extent of parental involvement in school activities such as curriculum committee participation, classroom volunteer participation, fund raising, or interacting with staff. A high score on this factor would mean that the principal views parental involvement with the school as high and positive in the many areas of school activities and operations.

The fourth factor, *Student Status*, consisted of four items related to the student characteristics in the school: the number of full-time students, the proportion of students living close to the school, and their access to computers. This factor could be considered an index of school size and proximity to students' home community. A high score would mean that the principal perceives the school to be of large size with a good level of computing resources, and with most, if not all, students living in the school catchment area.

Table 3: Five Factor Solution: School Questionnaire

Factor Label *Item stem (paraphrased)*	Questionnaire item #	Loading
Limitations to Learning (What percentage? 1= <,10%...3= >.25%)		
% of students with learning problems needing special attention	Q07D	0.623
% of students from single parent families	Q07E	0.485
% of students with health problems that inhibit learning	Q07F	0.684
(Asked to estimate extent of limiting capacity 1= None … 4 = A lot)		
School instructional capacity limited by lack of parental support	Q16A	0.688
School instructional capacity limited by student abilities	Q16B	0.735
School instructional capacity limited by student's home background	Q16C	0.821
School instructional capacity limited by community conditions	Q16D	0.744
Instruction limited by shortage of math teachers	Q17B	0.333
Climate (1=Strongly disagree … 4= Strongly agree)		
This school is supported by the community	Q31G	-0.601
Staff morale is high in this school	Q31H	-0.771
There is a strong school spirit	Q31I	-0.821
Students and staff take pride in this school	Q31J	-0.787
Parents (1=None … 4= A lot)		
Extent that parents act as volunteers in classrooms	Q28A	0.391
Extent that parents volunteer in monitoring student behaviour	Q28B	0.404
Extent that parents serve on committees on curriculum	Q28C	0.675
Extent that parents serve on committees on finance or administration	Q28D	0.671
Extent that parents influence selection of principal or teachers	Q28E	0.555
Extent that parents serve on committees on student conduct	Q28F	0.640
Extent that parents interact with staff on matters of their own child	Q28G	0.358
Extent that parents help raise funds for school	Q28H	0.435
Student Status (1=low … 4=high)		
Full-time equivalent students in school	Q05	0.776
Percentage of students within walking distance of school	Q07A	0.394
% of students using subsidized transportation	Q07B	-0.369
How many computers in school & available to students?	Q18	0.773
Factors of Student Achievement (1=Strongly disagree … 4= Strongly agree)		
Students can achieve if they work hard	Q31B	0.655
Students can achieve if they are taught well	Q31C	0.679
Students can achieve with adaptations for special needs	Q31D	0.525

The last factor, *Factors of Student Achievement*, consisted of three items asking principals about perceptions of correlates of student achievement: effort, good instruction, and adaptations for special needs. A high score on this factor means that the school principal views hard work by students, effective teaching, and school adaptations as major contributors to student achievement in mathematics.

Factor scores on each of the five factors were calculated for each school in the data set and appended to the school's record. This resulted in a set of 10 variables at level-2: five derived from the aggregated student factors and five from the school questionnaire.

MODELING STUDENT ACHIEVEMENT IN MATHEMATICS

A hierarchical linear modeling approach (Raudenbush & Bryk, 2002) was used in these analyses. The outcome measures as developed by the CMEC were the Content and Problem Solving tests of the SAIP mathematics assessment. The student-level correlates for the initial set of analyses reported in this paper consisted of the five factors scores at the student level (Table 2) and student gender (males coded 0 and females coded 1). The 10 school-level correlates were the five factors derived from the school questionnaires items (Table 3), and the aggregated five factor scores derived from the student questionnaire. As indicated earlier, the data were divided into four sets: by age (13 and 16 year-olds) and mathematics domain (Content and Problem Solving), resulting in four HLM models.

Explained Variance

The first step in the analyses was to evaluate the proportion of total variance in mathematics achievement scores that was accounted for by school level variables: the intra-class correlation or ρ (Raudenbush & Bryk, 2002). The proportion of variance in achievement scores in both domains attributable to schools (Table 4) was relatively low (<.20). There were slight differences across age groups with 13 year-old students having greater proportion of variance accounted for by schools (0.18 and 0.19) than that of the 16 year-old students (0.15 and 0.15).

Table 4: The proportion of variance in student achievement scores accounted for by schools

Test	ρ
Problem solving – age 13	0.18
Problem solving – age 16	0.15
Math content – age 13	0.19
Math content – age 16	0.15

Level-1 – Student Correlates of Mathematics Achievement. Next the student level predictors of mathematics achievement were modeled (Table 5). Two student factors were related to mathematics achievement for both ages and domains. The other three factors had weaker relationships that varied across age and domain.

For all four datasets, *Student Beliefs about Mathematics* was significantly related to mathematics achievement (>0.30). As positive views on the value of and interest in mathematics increased, mathematics achievement increased for students in both age groups and for both mathematics domains. The direction of influence is open to conjecture because positive views about mathematics could enhance interest and therefore achievement, but it could also be argued that high achievement stimulates positive views about mathematics. Likewise *Instructional Supports used by Students* was significantly related to mathematics achievement for all four data sets, but in an inverse manner (<-0.15). As reported use of instructional support increased, mathematics achievement tended to decrease. This may appear to be somewhat counterintuitive because parental assistance with homework is often thought to lead to an increase in student achievement. The response distributions for these items is highly skewed showing that the majority of students reported little or no use of any of the listed supports that may have influenced the nature of this relationship. However, because this is a student-reported perception, it may be that the students reporting

higher levels of use of supports tend to be weaker students, and this is consistent with other reported research (Ho & Willms, 1996).

Table 5: Student and School Level Coefficients

Correlate	CONTENT 13	CONTENT 16	PROBLEM 13	PROBLEM 16
Student-level				
Student math beliefs	.36	.33	.38	.37
Instructional supports	-.18	-.22	-.22	-.29
Instructional practices	.03	.08	.04	.10
Causes of math	-.08	0	-.06	0
Discipline climate	0	0	0	0
Gender	0	-.09	.10	.07
School-level correlates to intercept				
Limits to learning	-.14	-.21	-.14	-.18
Instructional supports	-.12	-.19	-.10	-.19
Causes of math	0	-.22	0	-.22
Discipline climate	0	0	0	-.17
Student math beliefs	.13	0	.11	0
School climate	0	0	-.04	-.05
Parent engagement	0	.05	0	.06
Student status	.07	.06	.05	0
Student achievement	.05	0	.05	0
Instructional practices	.08	0	0	0

a. Only coefficients significant at the $\alpha = 0.05$ level are reported, otherwise 0.

Instructional Practices was positively but weakly related to mathematics achievement for both age groups and domains of mathematics achievement. This result indicates that an increased diversity of classroom practices is associated with higher levels of mathematics achievement. A weak negative relationship (<-0.10) occurred between *Causes of Mathematics Performance* and mathematics achievement in both domains for 13 year-old students. As the strength of belief in these causes (both internal and external) increased, mathematics achievement for 13 year-olds tended to decrease. In contrast, no

significant relationship occurred between *Causes of Mathematics Performance* and achievement in the Problem mathematics domain for 16 year-old students.

Perceptions of classroom disorder and disruption in the mathematics classroom were not significantly related to mathematics achievement. Because this factor is based upon student perception elicited in two questionnaire items, it may well be that students are not particularly sensitive to evaluating disciplinary climate when based upon time estimates of classroom activities (Ma & Willms, 2004). More objective measures of classroom order and organization may shed light on the accuracy and consistency of student perceptions of classroom order and noise.

The relationships between student gender and mathematics achievement were weak and mixed. In the mathematics Content domain, 16 year-old males tended to do better than females although there was no significant gradient for 13 year-old students. However, in the mathematics Problem domain, females tended to have higher achievement than males in both age groups.

Level-2 – The Intercept: Mean School Mathematics Scores. The mean school mathematics scores (the intercept terms in the Level-1 models) were modeled using the 10 school-level correlates (Table 5). Even after conditioning on the school correlates, a significant variation occurred between schools in terms of the mean school mathematics achievement scores for both age groups and for both mathematics domains. This variation suggests that additional correlates of learning account for differences in mathematics achievement of students in the school that have not been included in these data sets.

Limitations to Learning and *Instructional Supports* were significant correlates of average school achievement in both mathematics domains and for both age groups. With the *Limits to Learning* factor, as principals perceived that limitations increase, the school mean mathematics scores tend to decrease. Not only was the direction of this relationship the same (negative) for all four datasets, the coefficients were very similar for each age. To an extent this negative relationship may be a self-fulfilling prophecy. Schools in which the principal views the students, their homes, and communities as problematic, thereby imposing limits on

what the school can do, tend to have lower achievement. In contrast, schools in which the principal views students and their homes and communities in a more positive light – with these elements not creating impediments to student learning – tend to be associated with better overall achievement. Alternatively, in schools with weak mathematics achievement the principal tends to develop a negative view, while in schools with strong mathematics achievement the principal tends to develop a positive view. It should be noted that these perceptions would likely be rather global impressions, for example, responses to the item related to the perceptions regarding single parent families are likely based on general impression rather than empirical evidence.

Similar to the results for Level-1, the *Instructional Supports* factor (averaged across students within each school) was negatively related to school mathematics performance (-0.10 to -0.19). This result means that as the average reported use of instructional supports within the school increases, the mean mathematics scores tend to decrease. Schools in which students report less use of instructional supports tend to have higher mathematics scores on average.

Causes of Mathematics was negatively related to mathematics performance for both Content and Problem Solving for the 16 year-old students (-0.22), but there was no significant relationship for the 13 year-old students. This result can be interpreted to mean that as 16 year-old students in the school increased their attribution of mathematics achievement to hard work and teacher and parent encouragement, the mean mathematics score for the school tended to go down. To an extent this relationship could be viewed as attributing performance in mathematics to external elements in the students' world: teachers and parents. However the inclusion of *hard work* in this factor, an internal element, is somewhat counterintuitive. This factor will need to be further investigated to better understand this relationship.

Disciplinary Climate is significantly predictive only for 16 year-old students in the Problem domain (-0.17). Here, as disciplinary climate as reported by students worsens (more disruption in mathematics classes, and higher scores on this factor) the mean school Problem achievement scores tend to decrease, a finding that is consistent with expectation. The absence of this relationship in the Content domain for both age groups,

and for 13 year-old students in the Problem domain, requires further investigation.

Student Mathematics Beliefs were positively related to average mathematics scores only for 13 year-old students. As the positive views about mathematics of students in the school increased, there tended to be a increase in the mathematics achievement of the students. This relationship does not seem to be present in the perceptions of 16 year-old students.

Level-2 – Modeling the Gradients. Gradients within the context of HLM5 (Raudenbush, Bryk, Cheong, & Congdon, 2000) are indices of the strength and direction of relationship between the student level correlates (e.g., *Mathematics Beliefs or Gender*) and the outcome measure (mathematics scores). The student gradients expressed as coefficients of the potential explanatory variables at Level-1 (Table 5) were modeled at the school level with Level-2 correlates.

The results for 13 year-olds in the level-2 model are shown in Table 6. Mathematics achievement in the Content and Problem domains for the 13 year-old students was influenced by their *Beliefs about Mathematics* (0.36 and 0.38) and their *Use of Instructional Supports* (-0.18 and -0.18). However, for the Content domain (Table 6) there were no significant Level-2 school correlates related to these two student gradients. For the Problem domain, the significant positive relationship between *Student Beliefs about Mathematics* and mathematics achievement (0.38) was influenced by two school-level correlates: aggregated or school average *Student Beliefs about Mathematics* (0.07) and by *Instructional Practices* (0.07). Hence for Problem Solving, the relationship between student beliefs about mathematics and achievement was not consistent across schools. Those schools with higher average scores on *Student Beliefs about Mathematics* and diversity of *Instructional Practices* had modestly higher Level-1 gradients between individual S*tudent Belief about Mathematics* and achievement. All other student gradients (*Instructional Supports, Instructional practices* and *Gender*) were essentially constant across schools and had no significant relationships to school-level correlates.

Table 6: School Level (Level-2) Coefficients: The Gradients

Student Variable Gradient (13yr/16yr)	School Correlates 13 year old	16 year old
Content Domain		
Student math beliefs (.36/.33)[a]	ns[c]	ns
Instructional supports (-.18/-.22)[a]	ns	Limits(-.03)
		Support(.15)
		Student Ach(.04)
Instructional practices (.03/.08)[a]	ns	ns
Causes of math (-.08/-.04)[a]	ns	ns
Gender (0/-.09)	ns	ns
Problem Solving		
Student math beliefs (.38/.37)[b]	Student Beliefs(.07)	Parents(-.03)
	Instruction Practices(.07)	Discipline(-.09)
Instructional supports (-.22/-.29)[a]	ns	ns
Instructional practices (.04/.10)[a]	ns	ns
Gender (.10/.07)[a]	ns	ns

a. No significant variation (τ) between schools on this gradient
b. Significant variation (τ) between schools on this gradient
c. No significant school-level correlates

For 16 year-old students in the Content domain (Table 6), only the *Instructional Supports* gradient varied significantly across schools. As the average level of student *Use of Instructional Support* within a school (the Level-2 correlate – the school average) increases, the influence of student *use of instructional* supports (the Level-1 gradient) flattens. In other words the relationship of student use of supports to mathematics achievement

which is negative (-.22) is reduced. In addition, two other school correlates are significantly related to this gradient: *Perceived Limitations to Learning* (-.03) and *Factors of Student Achievement* (.04). However, the values are quite small, suggesting that the relationship of school correlates to student-level correlates is weak.

In the Problem domain, only one Level-1 gradient, *Student Beliefs about Mathematics,* varied significantly across the schools in the 16 year-old student sample (Table 6). And this student correlate is negatively influenced by average parental involvement in schools (*Parents*) and by the average level of disruption and noise in the mathematics classrooms (*Discipline*). However, the values (-0.03 and –0.09, respectively) are small, suggesting that the relationship to student performance as measured by mathematics achievement is slight. The relationship of average student perception of classroom disruption (*Discipline*) certainly is in line with expectations: increased disruption related to a decrease in the positive relationship of *Student Beliefs about Mathematics* on mathematics achievement. However, the weak negative relationship of average reported levels of parental involvement with school affairs (*Parents*) is counterintuitive and requires further investigation.

DISCUSSION

The results reported in this article reveal some consistent strong relationships at the student level and an important characteristic of large-scale assessment data in Canada. Students' beliefs about mathematics are positively related to mathematics achievement as measured by the SAIP test of Content and Problem Solving for both age groups and for both domains of mathematics. Also the perceptions of school principals of the school environment (*Limitations to Learning*) were related to overall school performance. We see that there is a relationship between perceptions or attitudes and mathematics achievement at both the student and school levels. The policy relevance here is clear: enhancing student and principal perceptions and beliefs could enhance student achievement and therefore school performance.

Second and somewhat counterintuitively, as students report that the use of instructional supports (such as parental assistance with mathematics homework and the use of computers in the mathematics

classroom) increases, there is an associated decrease in mathematics scores. The policy relevance here is unclear because to reduce parental involvement with a students' work at home on school subjects would not be accepted by most educators and parents. As indicated earlier, inspection of the response distributions for the items related to instructional support revealed that the majority of students reported little or no use of any of the supports listed in the student questionnaire. It may be that the students who did report they made use of support were low performing students and thus the negative relationship. It may also be that this factor is an index of student independence: those students reporting less use of parental assistance are relatively independent in their work on mathematics and this is associated with higher levels of performance. However, from a research perspective the direction is clear, the need exists for collecting more direct indicators of instructional support use to better understand its function in relation to student achievement in a general manner. It will be of interest to track these patterns through similar future analyses conducted on subsets of the data for different educational jurisdictions.

A pattern that emerged in these analyses of SAIP Mathematics 2001 was the general stability of these student-level relationships (the gradients) across schools. Most of the student-level coefficients did not show significant school variation and, as a consequence, many did not have any significant school-level correlates influencing their values, suggesting that the correlates have a common influence in schools nationally.

In regard to large-scale assessment data, the analysis of the four datasets from SAIP Mathematics 2001 showed that less than 20 per cent of variation in achievement measures was attributable to schools. This is a characteristic and significant feature of large-scale assessment data in Canada (OECD, 2001; Anderson, 2002): the relatively low proportion of variation in student achievement that can be attributed to schools (<20%). An international average of OECD countries as assessed in the *Programme of International Student Assessment (PISA)* is .35 (OECD, 2001). Although in Canada school differences are modest, the influence of these differences on overall student achievement should certainly be considered by those who publish school rankings (e.g., Cowley, Easton,

& Walker, 2001) because these rankings imply that most if not all variation in averaged school results, regardless of how they are derived, are attributable to school-based factors. The evidence does not support this implication.

In the results reported here it is clearly evident that mean school performance did vary significantly across schools, and these scores were related to school characteristics, in particular the extent to which principals reported perceiving student, home, and community traits as limiting elements of school effectiveness, and the average reported student use of instructional supports. This first relationship does appear to be accessible to policy intervention: principals' beliefs could be influenced by communications strategies in the school system. The second relationship is similar to that found at the student level for the use of instructional support: the more use of or dependence upon instructional support reported by students, on average within the school, the lower the associated mathematics scores. There is clearly a relationship here, but it would be important to collect further data on these variables to ensure that what students are reporting is indeed an accurate reflection of the extent to which they do homework and the extent to which parents assist them, and to collect direct evidence of the instructional practices used in mathematics such as computer use, the use of reference materials, or resource people for mathematics classes. This should help to better understand these relationships and their utility in policy formulation and implementation.

This introduces an important direction these analyses identify in evaluating the utility of the SAIP datasets for statistical modeling of school performance: the need for more and better data on student, school, class, and home variables, an issue more fully discussed in Rogers et al. (2006). The negative relationship between use of instructional supports to mathematics achievement, the non-significant relationship between classroom disruption and noise to mathematics achievement, and the startling stability of student gradients across schools all suggest the need for further investigation with better, more direct measures of correlates of learning outcomes and the development of instrumentation to collect this information. An obvious need exists to collect data at the class or teacher level because less than 20 per cent of

variance in student performance can be attributed to schools, meaning that over 80 per cent can be attributed to students and classes. The analyses reported here were based upon the results generated by SAIP Mathematics 2001, in effect statistical modeling based upon found data. The findings of studies such as this should be used to inform the shape of future assessment programs. Because data such as these are collected annually in provincial and national assessment programs, with appropriate planning and design, data collection strategies could be modified with little or no additional expenditure to collect data to better address policy- and practice-relevant evidence requirements.

NOTES

[1] The COLO project is funded by the Canadian Educational Statistics Council and the Social Sciences and Humanities Research Council.

REFERENCES

Anderson, J. O. (2002). School community characteristics and student achievement in grade 10 reading, writing and numeracy of the British Columbia Foundational Skills Assessment. In L. Hayduk, X. Ma, & C. Carter-Snell (Eds.), *Structural equation modeling and hierarchical linear modeling: Communicating across disciplines* (pp. 139-160). Edmonton, AB: Population Research Laboratory , University of Alberta.

Council of Ministers of Education Canada. (2001). Report on mathematics assessment III: School Achievement Indicators Program 2001. Toronto: Council of Ministers of E Canada.

Council of Ministers of Education Canada. (2005). SAIP Mathematics III 2001. Retrieved September 18, 2006, from http://www.cmec.ca/saip/math2001/indexe.stm

Cowley, P., Easton, S., & Walker, M. (1999). The 1999 report card on British Columbia's secondary schools. *Public Policy Sources, Number 22.* Vancouver, BC: Fraser Institute.

Dumais, S. A. (2002). Cultural capital, gender, and school success: The role of habitus. *Sociology of Education*, 75(January), 44-68.

Education Quality & Accountability Office (2002). *EQAO Provincial Achievement Results 2001*-2001. Toronto: The Author.

Fitz-Gibbon, C. (1998). Indicator systems for schools: Fire fighting it is! In D. Shorrocks-Taylor (Ed.), *Directions in Educational Psychology* (pp. 189-205). London, UK: Whurr Publications Ltd.

Fuchs, T., & Woessmann, L. (2004). *What accounts for international differences in student performance? A re-examination using PISA data*. Munich, Germany: University of Munich, Ifo Institute for Economic Research.

Hanushek, E. A. (2002). The seeds of growth. *Education Next*, Fall 2002. Retrieved Septembr 18, 2006, from http://www.educationnext.org/20023/10.html

Hattie, J., Jaeger, R. M., & Bond, L. (1999). Persistent methodological questions in educational testing. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of Research in Education Volume* 24 (pp. 393-446). Washington, DC: American Educational Research Association.

Hill, P. W., & Crévola, C.A. (1999). Key features of a whole-school, design approach to literacy teaching in schools. *Australian Journal of Learning Disabilities*, 4(3), 5-11.

Ho, E. S., & Willms, J. D. (1996). Effects of parental involvement on eighth-grade achievement. *Sociology of Education*, 69, 126-141.

Joint Advisory Committee. (1993). *Principles for fair student assessment practices for education in Canada*. Edmonton, AB: Centre for Research in Applied Measurement and Evaluation (CRAME), University of Alberta.

Keeves, J.P. (2002, October). Why do education systems differ in educational outcomes? A paper presented at the *Providing world-class school education,* Australian Council for Educational Research Conference. Sydney, NSW, Australia.

Kennedy, M. M. (1999). Approximations to indicators of student outcomes. *Educational Evaluation and Policy Analysis*, 21(4), 345-363.

Ma, X. (2001). Stability of school academic performance across subject areas. *Journal of Educational Measurement*, 38(1), 1-18.

Ma, X., & Willms, J. D. (2004). School disciplinary climate: Characteristics and effects on eighth grade achievement. *Alberta Journal of Educational Research*, 50(2), 169-188.

Mandeville, G. K., & Anderson, L.W. (1987). The stability of school effectiveness indices across grade levels and subject areas. *Journal of Educational Measurement*, 24(3), 203-216.

Mulaik, S. A. (1972). *The foundations of factor analysis*. New York: McGraw-Hill.

Organization for Economic Cooperation and Development. (2001). *Knowledge and skills for life: First results from the OECD programme for international student assessment (PISA) 2000.* Paris: The Author.

Raudenbush, S. W. (2002). Scientifically based research. A paper presented at the Seminar on Scientifically Based Research - US Department of Education. Retrieved September, 2006, from: http://www.ed.gov/nclb/methods/whatworks/research/page_pg5.html

Raudenbush, S. W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods (*2nd Ed*)*. Newbury Park, CA: Sage Publications.

Raudenbush, S. W,, Bryk, A. S., Cheong, Y. F., & Congdon, R. T. (2000). *HLM5: Hierarchical linear and nonlinear modeling.* Chicago: Scientific Software International.

Raudenbush, S.W., & Willms, J. D. (1991). The organization of schooling and its methodological implications. In S. W. Raudenbush & J. D. Willms (Eds.), *Schools, classrooms and pupils: International studies of schooling from a multilevel perspective (pp. 1-12).* New York: Academic Press, Inc.

Reeves, E. B., & Bylund, R. A. (2005). Are rural schools inferior to urban schools? A multilevel analysis of school accountability trends in Kentucky. *Rural Sociology*, 70(3), 360-386.

Rogers, W. T., Anderson, J. O., Klinger, D. A., Dawber, T., Feng, A. & Zhou, J. (2006). Pitfalls and potential of secondary data analysis of the Council of Ministers of Education, Canada National Assessments. *Canadian Journal of Education 29(3)*, 757-770.

Rogers, W. T., Ma, X., Klinger, D. A., Dawber, T., Hellsten, L., Nowicki, D., & Tomkowicz, J. (2006). Examination of the influence of selected factors on performance on Alberta Learning Achievement Tests. *Canadian Journal of Education, 29(3)*, 731-756.

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal*, 32(4), 583-625.

Shore, A., Madaus, G., & Clarke, M. (2000). Guidelines for policy research on educational testing. *NBETPP Policy Statements 4(1).* Boston: National Board on Educational Testing & Public Policy Statements. Retrieved September, 2006, from: http://escholarship.bc.edu/lynch_facp/66/)

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, 31(7), 15-21.

Statistics Canada & Council of Ministers of Education Canada, (2000). *Education indicators in Canada: Report of the Pan-Canadian indicators program 1999*. Ottawa & Toronto: The Authors

Systat (2002). *Systat 10.2: Statistics I.* Richmond, CA: Systat Software Corp.

Watt, T. T. (2003). Are small schools and private schools better for adolescents' emotional adjustment? *Sociology of Education*, 76(4), 344-367.

Willms, J. D. (2004). *Reading achievement in Canada and the United States: Findings from the OECD programme for International student assessment.* Ottawa, ON: Human Resources Skills and Development Canada Publications.

Willms. J. D., & Kerckhoff, A. C. (1995). The challenge of developing new educational indicators. *Educational Evaluation and Policy Analysis*, 17(1), 113-131.

Willms, J. D., & Somers, M-A, (2001). Family, classroom, and school effects on children's educational outcomes in Latin America. *School Effectiveness and School Improvement*, 12(4), 409-445.